

teaches modifying a speaker recognition (verification/identification) system. Speech recognition is defined in the McGraw-Hill Dictionary of Scientific and technical terms as:

“The process of analyzing an acoustic speech signal to identify the linguistic message that was intended, so that a machine can correctly respond to spoken commands.”

In contrast, the TechEncyclopedia found at www.techweb.com defines speaker recognition as:

“The ability to recognize a person by his or her spoken voice. This is used for security purposes, not voice recognition. Like voice recognition however, the user is required to train the system by speaking certain phrases. Contrast with voice recognition. See biometrics.”

A speech recognition device identifies what is spoken. A speaker recognition device identifies who is speaking.

Within the Office Action, it is stated that Kanevsky does teach a system to modify a speech recognition system. To support this assertion, column 8, lines 48-50 is cited. Specifically, column 8, lines 48-50 of Kanevsky states:

“However, the system collects voice samples from the caller’s answers to the plurality of questions and builds a user voice model (e.g., user model 20) therefrom.”

As can be seen from the passage above, there is no indication of a speech recognition system (ASR 28). Column 8, lines 48-50 of Kanevsky simply refers to building a user model 20. To correctly interpret this passage, the surrounding passages must also be considered. Specifically, column 8, lines 37-48 states:

“...if a caller calls the central server 22 for the first time and the system has a database information pertaining to the caller but does not have an acoustic model set up for that caller, the following procedure may be performed. The central server 22 asks a plurality of questions from the database, the number of questions depending upon the known

average error rate associated with ASR 28 and semantic analyzer 40. Then, based only on the scores achieved by the answers received to the questions, the server 22 makes a determination whether or not to permit access to the caller.”

In other words, the user model 20 is built by asking questions and receiving answers, where the answers are then collected by the system to build the user model 20. Further, column 8, lines 51-55 states:

“Accordingly, the next time the caller calls, the server 22 need ask only a few random questions from the database and, in addition, use the text-independent speaker recognition module 52 along with the new user model to verify his identity, as explained above.”

In other words, the text-independent speaker recognition module 52 uses the user model 20 to verify a user identity. **At no point does Kanevsky teach that the ASR 28 uses the user model 20.** As such, modifying the user model 20 with additional voice samples has no impact on the performance of the ASR 28, as the ASR 28 does not use the user model 20 to perform speech recognition.

Within the Office Action, it is stated that Kanevsky teaches “to modify a speech recognition system from speech model (answer of questions to identify the speaker) to speaker dependent model (User model).” However, there is no support within Kanevsky to substantiate this conclusion. Specifically, **modification of the user model 20 is not modification of a speech recognition system.** As described above, the user model 20 is distinct from the ASR 28. The only interaction between the ASR 28 and the user model 20 is a one-way data flow **from the ASR 28 to the user model 20** when voice samples are first recognized by the ASR 28 and then sent to the user model 20 to build the user voice model. **The ASR 28 does not use the user model 20 to perform the speech recognition.** In fact, Kanevsky clearly teaches that the user model 20 is used by the text-independent speaker recognition module 52 and a similar voice classification module 68, and not the ASR 28. Specifically, column 10, lines 54-60 of Kanevsky states:

“As previously explained, a user model is employed to estimate a probability of a

particular user's identity. The model is utilized in accordance with the text-independent speaker recognition module 52 (FIG. 2) and the voice classification module 68 (FIG. 3)."

From the discussion above, it is clear that the user model 20 is modified by voice samples processed by the ASR 28 which are provided to the user model 20 from the ASR 28. It is also clear that the ASR 28 does not use the user model 20 to perform its speech recognition processing. Therefore, at no point in Kanevsky are the collected voice samples of the user model 20 used to modify the ASR 28, as proposed within the Office Action.

In fact, it is recognized within Kanevsky that the ASR 28 includes functional limitations. To account for these limitations, the system of Kanevsky does not teach a process of modifying the ASR 28, but instead it is preferred that more than one question be asked prior to making a decision to permit or deny access to the speaker to account for possible recognition or understanding errors associated with the ASR 28 (Kanevsky, col. 6, lines 44-56). Clearly, if the ASR 28 was designed to be modified, this would be stated instead of the above passage.

Further, Kanevsky acknowledges the use of a conventional speech recognition device (Kanevsky, col. 13, lines 50-61). As such, since Kanevsky teaches a conventional ASR 28, and the ASR 28 is not modified within the speaker identification system of Kanevsky, Kanevsky does not teach the use of a speaker-specific **modified** speech recognition system as claimed in the independent Claims 1, 17, 32, 48, and 54 of the present invention.

In contrast to the teachings of Kanevsky, the speech recognition system of the present invention provides for speaker-specific acoustic models to be used in the **speech recognition process**. Multiple users can access the same application, each user having an individualized speaker-specific modified speech recognition system which is stored, retrieved from storage, and modified according to samples of the specific speaker's speech. Kanevsky teaches a system to modify a **speaker identification process**. Kanevsky does not teach a system to modify a **speech recognition system**.

Within the Office Action, it is stated that the present invention teaches a method to modify a speaker identification system. The Applicant respectfully disagrees. To support this assertion, it is stated within the Office Action that the independent claims of the present invention claim "obtaining an identification of a speaker by the speaker's name" (emphasis added). There is no such limitation claimed by the Applicant. The Applicant respectfully points

out that the independent claims, and specifically independent Claim 1, teaches “obtaining an identification of a speaker.” There is no limitation as to how the identification is obtained. In fact, the specification states that any number of conventional identification techniques can be used including prompting the speaker to speak his name, entering a personal identification number, entering an account number, automatically receiving the speaker’s caller ID for a telephone call, and utilizing voice identification techniques (Specification, page 5, lines 4-9).

The independent Claim 1 is directed to a method of adapting a speech recognition system. The method of Claim 1 includes the steps of obtaining an identification of a speaker, obtaining a sample of a speaker's speech during a first remote session, recognizing the speaker's speech utilizing the speech recognition system during the first remote session, modifying the speech recognition system according to the sample thereby forming a speaker-specific modified speech recognition system, storing a representation of the speaker-specific modified speech recognition system in association with the identification of the speaker, and using the representation of the speaker-specific modified speech recognition system to recognize speech during a subsequent remote session with the speaker. As discussed above, Kanevsky teaches a system to modify a speaker identification process. Kanevsky does not teach a system to modify a speech recognition system, and Kanevsky does not teach storing and re-using the modified speaker-specific speech recognition system. For at least these reasons, Claim 1 is allowable over the teachings of Kanevsky.

Claims 2-10 and 12-16 are each dependent upon the independent Claim 1. As discussed above, the independent Claim 1 is allowable over the teachings of Kanevsky. Accordingly, Claims 2-10 and 12-16 are each also allowable as being dependent upon an allowable base claim.

Within the Office Action, Claims 17-26, 28-40, and 42-58 have been rejected as having similar limitations as Claims 1-10 and 12-16. The Applicant respectfully traverses this rejection for at least the same reasons as discussed above pertaining to Claims 1-10 and 12-16.

Rejections Under 35 U.S.C. § 103

Claims 1-10, 12-26, 28-40, and 42-58 stand rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent No. 5,127,055 issued to Larkey. The Applicant respectfully traverses this rejection.

Larkey teaches the digitizing, processing, and analyzing of incoming speech and

comparing the incoming speech to reference patterns stored in a reference pattern storage memory (col. 4, lines 10-16). A data processing system then makes a best estimate of the identity of the incoming signal and provides electrical signals identifying the best estimate to an output device (col. 4, lines 16-20). To clarify, "identify" does not refer to the identification of the user; instead, "identify" refers to the recognizing of the incoming user utterance. The stored reference patterns are dynamically updated and adapted by using correction actions the user has provided about the correctness of the recognition, the correction actions being critical to successful operation of the reference pattern adaptation method (col. 4, lines 32-50).

The independent Claims 1, 17 and 32 as amended in the "Amendment and Response to Office Action Mailed on October 20, 2000", which was mailed by the Applicant on January 22, 2001, and the independent Claims 48 and 54 as amended in the "Amendment and Response to Final Office Action Mailed on April 11, 2001", which was mailed by the Applicant on June 11, 2001, all include the limitation of obtaining an identification of a user. Larkey does not teach obtaining an identification of a user.

Further, since Larkey does not teach obtaining an identification of a user, Larkey can not possibly teach the limitation of storing a speaker-specific modified speech recognition system in association with the identification of the user, as claimed in the independent Claims 1, 17, 32, 48, and 54. The speech recognition system of the present invention not only identifies the user but uses that identification to modify and store a speaker-specific acoustic model that correlates directly with the identified user.

Within the Office Action, it is acknowledged that Larkey does not explicitly teach a "modified speech recognition system in association with an identification of the speaker." However, within the Office Action, column 1, lines 51-56 of Larkey is cited which teaches "a dynamic reference pattern updating mechanism for improving the precision with which incoming unknown speech can be identified, and providing reference patterns which better characterize a speaker's manner of pronouncing a selected word vocabulary." According to the Office Action, this cited passage of Larkey indicates that it would have been obvious to one of ordinary skill in the art to "train speaker [recognition system] by knowing speaker's manner of pronouncing a selected word vocabulary as to recognize speech and user at the same time." The Applicant respectfully disagrees with this conclusion. The cited passage clearly indicates that "incoming unknown speech can be identified", not that an incoming **speaker** is identified. As previously

discussed, “identify” does not refer to the identification of the user; instead, “identify” refers to recognizing an incoming user utterance. Further support for this use of the word “identify” is provided in column 4, lines 13-20 of Larkey, which states:

“The data processing unit digitizes, processes and analyzes the incoming speech and compares the incoming speech to reference patterns stored in a reference pattern storage memory 16. The data processing system then makes a best estimate of the identity of the incoming speech and provides electrical signals identifying the best estimate to an output device such as a display terminal 18.”

As such, the speech recognition system of Larkey does not teach recognizing the user as proposed within the Office Action.

Further, Larkey unambiguously teaches a speech recognition apparatus, not a speaker recognition apparatus. As discussed above, it is well known in the art that speaker recognition is distinct from speech recognition. Since a speech recognition device does not perform the same function as a speaker recognition device, and in absence of any teachings to the contrary, it is not obvious that the speech recognition device of Larkey can recognize speech and can recognize a user at the same time, as proposed within the Office Action.

The independent Claim 1 is directed to a method of adapting a speech recognition system. The method of Claim 1 includes the steps of obtaining an identification of a speaker and storing a representation of the speaker-specific modified speech recognition system in association with the identification of the speaker. As discussed above, Larkey does not teach obtaining an identification of a user. Larkey also does not teach storing a representation of a speaker-specific modified speech recognition system in association with the identification of the speaker. For at least these reasons, the independent Claim 1 is allowable over Larkey.

Claims 2-10 and 12-16 are each dependent upon the independent Claim 1. As discussed above, the independent Claim 1 is allowable over the teachings of Larkey. Accordingly, Claims 2-10 and 12-16 are each also allowable as being dependent upon an allowable base claim.

Within the Office Action, Claims 17-26, 28-40, and 42-58 have been rejected as having similar limitations as Claims 1-10 and 12-16. The Applicant respectfully traverses this rejection for at least the same reasons as discussed above pertaining to Claims 1-10 and 12-16.

For the reasons given above, Applicant respectfully submits that all of the remaining claims are in a condition for allowance, and allowance at an early date would be appreciated. Should the Examiner have any questions or comments, he is encouraged to call the undersigned at (408) 530-9700 to discuss the same so that any outstanding issues can be expeditiously resolved.

Respectfully submitted,
HAVERSTOCK & OWENS LLP

Dated: December 24, 2002

By: Jonathan O. Owens
Jonathan O. Owens
Reg. No. 37,902
Attorneys for Applicants

CERTIFICATE OF MAILING (37 CFR § 1.8(a),
hereby certify that this paper (along with any
referred to as being attached or enclosed) is being
deposited with the U.S. Postal Service on the date
shown below with sufficient postage as first class
mail in an envelope addressed to the: Assistant
Commissioner for Patents, Washington D.C. 20231

HAVERSTOCK & OWENS LLP

Date: 12/24/02 By: File